Wireless Pers Commun (2013) 72:2185–2214
DOI 10.1007/s11277-013-1143-7

Star-Structure Network Coding for Multiple Unicast Sessions in Wireless Mesh Networks

Alireza Shafieinejad · Faramarz Hendessi · Faramarz Fekri

Published online: 3 April 2013

Abstract In this paper, first, we propose Star-NC, a new network coding (NC) scheme for multiple unicast sessions in an n-input n-output star structure. Then, we evaluate the network throughput of this coding scheme in wireless mesh network over the traditional non-NC transmission. Our scheme benefits from the proximity of all the nodes around the relay node and employs a more general form of overhearing different from other schemes such as COPE. We found that the gain of our NC scheme depends on both the star size and the routing pattern of the unicast transmissions. Based on this, we identify both the situations which the maximum gain is achievable and a lower bound for the expected value of the gain in the case of random routing pattern. Next, we propose an analytical framework for studying throughput gain of our Star-NC scheme in general wireless network topologies. Our theoretical formulation via linear programming provides a method for finding source-destination routes and utilizing the best choices of our NC scheme to maximize the throughput. Finally, we evaluate our model for various networks, traffic models and routing strategies over coding-oblivious routing. We also compare the throughput gain of our scheme with COPE-type NC scheme. We show that Star-NC exploits new coding opportunities different from COPE-type NC and thus can be used with or without this scheme. The results show that Star-NC has often better performance than COPE for a directional traffic model which is a typical model in wireless mesh networks. Moreover, we found that, joint Star and COPE-type NC has better throughput performance than each of Star or COPE alone.

A. Shafieinejad (✉) · F. Hendessi
Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran
e-mail: a.shafieinejad@ec.iut.ac.ir

F. Hendessi
e-mail: hendessi@cc.iut.ac.ir

F. Fekri
Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: faramarz.fekri@ece.gatech.edu

Keywords Network coding · Wireless mesh network · Unicast routing · Star topology · Grid network

1 Introduction

Wireless networks provide means for mobility, internet connectivity and distributed sensing. However, the throughput limitation of these networks imposes many practical problems. The NC scheme proposed in [1] has been shown to increase throughput of wireless networks by employing the broadcast nature of wireless media. In the work of Li et al. [2], it was shown that linear codes are sufficient to achieve maximum throughput for multicast traffic. Unicast, on the other hand, has received less attention relative to the multicast. In [3–6], specific topologies with unicast routing were studied and shown that network coding results better throughput than a flat routing. In [7], COPE was introduced, which is a packet encoding scheme via XOR operation. It studied certain basic topologies such as chain, cross and wheel in a bidirectional unicast traffic model and the throughput gain was reported by the first testbed deployment of wireless network coding.

In the follow up [8] extends the use of COPE in any wireless network with any patterns of multiple concurrent unicast transmissions. From a theoretical perspective, the authors provided two linear programming formulations for measuring the throughput improvements of COPE-type NC for both with and without opportunistic listening. By these formulations, the authors advocated the idea of *coding-aware routing*, i.e. the routing is made aware of NC opportunities.

Some of the work in the literature considers *pairwise network coding*, i.e., NC for two unicast sessions. The work in [9] is focused on butterfly topology (a well-known example for multicast NC [1]) and proposed a scheme similar to the one in [4]. The authors tried to exploit NC opportunities which originate from this coding scheme in a general wireless network. Another important work is [10] wherein the PINC-*Pairwise Intersession Network Coding* scheme is proposed. This scheme has a graph-theoretic approach and is based on the concept of the paths with *controlled edge-overlap* which is an extension of *edge-disjoint* paths for non-coding scheme. The work in [11] uses PINC and proposes an optimal joint coding, scheduling and rate-control scheme.

In [12] and [13], the authors studied the benefits of NC in multi-channel wireless networks. Both of them employ the simple COPE NC scheme without opportunistic listening known as Alice-Bob topology. In [12], in addition to network coding and routing, a randomized channel assignment algorithm was used together with the LP formulation. In [13], the formulation was provided for a new coding named as *analog NC* i.e. a simple physical-layer coding method mixes simultaneously arrived radio waves, in analog domain, instead of the conventional NC scheme. The authors also develop a heuristic joint link scheduling, channel assignment, and routing algorithm that aims at approaching the optimal solution to the optimization problem.

Further, a large portion of works has studied the limitation of network coding. In [14] the authors studied the throughput improvement of network coding in the Gupta and Kummar model [15]. They showed that the throughput of arbitrary NC schemes is improved by a constant order in the protocol model. They further derived the bounds for the throughput benefit ratio, i.e., the constant order throughput improvement using network coding in 1D, 2D and 3D spaces. But they did not provide any scheme that achieves the upper bounds. In [16], authors analyzed the performance of the coding scheme under realistic setting such as the physical layer characteristics and the random medium access. However, the main performance metric in [16] is limited to the encoding number, i.e., the number of packets that can be encoded by a node in each transmission.

Recently in [17] and [18], we used a simple form of the Star-NC for a 3-input 3-output star structure with a constant routing pattern in a multi-channel/interface network. The proposed scheme in [18] is based on the combination of the *coded-overhearing* and *coding-aware* channel assignment. The proposed algorithm overcome the radio coverage limitations in conventional NC schemes, and improves the aggregate throughput when there are an insufficient number of interfaces.

In this paper, we extend the concept of Star-NC in both the size of the star structure and the routing pattern of unicast transmissions. Our NC scheme considers multiple unicast flows that intersect each other at a relay node. The relay node decreases the number of transmissions by mixing the packets of these flows (sessions). This reduction is due to the opportunistic listening of the nodes at the proximity of the relay node. The key idea of our NC scheme is the generality and flexibility of the opportunistic listening which exploits more coding opportunities, thereby reducing more transmissions of the relay nodes. We focus on the star structure as a primitive element of network coding for the unicast traffic. Indeed, in an arbitrary network, each node with its neighbors creates a star structure. For the node with a larger degree, the probability by which some unicast sessions intersect each other at that node becomes higher and thus more opportunity for the Star-NC scheme will be created around the node. Since the star structure acts as a primary element of wireless networks, its analysis can provide a tool for understanding and optimizing the large and complex network topologies.

Note that in this paper, we do not define a full-fledged network coding protocol, but instead propose a novel structure for network coding which opens a new paradigm to achieve more throughput, namely how to maximize the amount of data delivered in a single transmission. We focus on algorithmic analysis (using LP based formulations) that determines the potential benefits of our NC scheme across arbitrary wireless topologies, traffic demands, as well as impact of joint network coding and interference-aware routing techniques. The main contributions of this paper can be summarized as follows:

- Under realistic assumptions about the radius of transmission and overhearing, we propose an NC scheme for multiple unicast sessions with a flexible and general form of opportunistic listening.
- Our scheme exploits more coding opportunities than other schemes (such as COPE) in directional traffic which is a typical form of traffic for wireless mesh networks that provided Internet connectivity.
- We derive an upper and a lower bound for the coding gain and the corresponding pattern of unicast transmissions which leads to these bounds. We further obtain a lower bound for the expected Star-NC gain when multiple unicast transmissions intersect each other with a random pattern.
- We provide a linear programming formulation to evaluate the Star-NC for any wireless network, any traffic model and any routing method. Our formulation allows the integration of our NC scheme with other coding schemes such as COPE.
- We consider the integration of multi-path routing with Star-NC schemes to increase the network throughput. Thus, the selection of the paths must be done with the awareness of both coding opportunities and interference among nodes, i.e. the *interference-aware* and *coding-aware* routing.

The rest of this article is organized as follows. In Sect. 2, we introduce the Star-NC, its benefits and also its limitation in practice. We analyze the Star-NC in Sect. 3 including the existence of the coding scheme, bounds and the expected value of the gain. We also propose an algorithm to find the proper coding scheme for any concurrent set of n-unicast sessions. In Sect. 4, we develop a theoretical formulation to study benefits of our coding scheme over non-coding routing and routing with the COPE-type coding scheme. In Sect. 5, using our formulation, we evaluate the benefit of our coding scheme with a variety of network topologies, traffic model and routing strategies. Section 6 introduces some of existing wireless mesh networks which are suitable for our NC scheme. Finally, Sect. 7 concludes the paper.

2 Star-NC: Overview and Motivating Examples

2.1 The concept

We focus on star structures as a basic element of network coding for multiple unicast sessions. A star structure of size n is composed of n input nodes, a relay node and n output nodes. The main goal is forwarding the packets from input nodes via the relay node to output nodes. In other words, the relay node is the cross point of n different paths corresponding to n unicast transmissions in the network. This structure acts as a switch in the network and forwards the incoming packets to next hop nodes as an intermediate step of routing. The basic idea was explained in Fig. 1a, b where the transmitted and overheard packets are indicated by solid and dotted lines, respectively. $S_1, S_2, \ldots$ and S_n denote the input nodes while $D_1, D_2, \ldots$ and D_n denote the output ones. Also, the relay node is denoted by M. Figure 1a shows a *full-star* structure of size 3 while Fig. 1b depicts a *partial-star* structure of size 2. We identify the full-star from partial-star with a concept of *overhearing direction*. It refers to the direction of overheard node to listener node respect to the relay node. In full-star structure, such as Fig. 1a, we can see overhearing of both clockwise and counter clockwise directions. However, in partial-star structure the overhearing is limited to be in one of directions, e.g. clockwise in Fig. 1b. This limitation can be originated from the geographical position of the nodes around the relay node. In full-star, the first output node is in the transmission range of the first input node. The same is true for the last output and input nodes. However, in the partial-star, only the first input and output nodes are in the transmission range of each other. Indeed, the last output node is far from any of the input nodes and, hence, does not have a direct access to any of them.

The goal is to route the incoming packets of input nodes to corresponding output nodes. In Fig. 1a, the incoming packets (P_1, P_2, P_3) must be routed in sequence to (D_2, D_3, D_1). Similarly in Fig. 1b the incoming packets (P_1, P_2) are to be routed to (D_2, D_1), respectively. These packets belong to different independent unicast sessions which include the paths $S_1 - M - D_2, S_2 - M - D_3, S_3 - M - D_1$ for Fig. 1a and $S_1 - M - D_2, S_2 - M - D_1$ for Fig. 1b, respectively. In standard routing, this is done by n unicast transmissions via the relay node M to output nodes. However, as depicted in the corresponding figures, Star-NC can do it by a single broadcast transmission, in an encoded packet, to the output nodes in both cases. In fact, two transmissions in Fig. 1a and one in Fig. 1b are saved for M as the relay node. Note that, in the general case, D_1 through D_n are not necessarily the final destination of P_1 through P_n. Instead, these nodes act as intermediate nodes, forwarding the packets to other nodes such as $\acute{D}_1$ through $\acute{D}_n$. In the rest of this article, for simplicity, the nodes $\acute{D}_1$ through $\acute{D}_n$ are omitted from figures.

Precisely, the network coding for this structure, here after referred to Star-NC, consists of three steps:

1. ***Coding at input nodes:*** S_1 through S_n send their encoded packets to the relay node.
2. ***Coding at relay node:*** Node M broadcasts its encoded packets (e.g. $P_1 + P_3$ in the above example) to output nodes.
3. ***Decoding at output nodes:*** D_1 through D_n decodes the inteneded packets using both encoded packets received from M and the overheard packets from neighbors (i.e. a subset of input and output nodes) and then forwards them to the next hops. For example, D_3 in Fig. 1a, decodes its packet (P_2) by means of XORing $P_1 + P_2 + P_3$ overheard from S_3 and the encoded packet $P_1 + P_3$.

Note that for some schemes, such as Fig. 1b, the first step is not needed since no coding is done by input nodes. It means that there is not any opportunistic listening for the input nodes. Thus the Star-NC without/with coding at input nodes will be a *two-hop/three-hop* coding scheme.

Each scenario of Star-NC corresponds to a specific routing pattern of flows between input and output nodes. Each pattern is identified by a unique and spanning mapping from input nodes to output nodes. We define *target permutation* of Star-NC as a permutation of $(1, 2, \ldots, n)$ that identifies the receiving packet indexes of the output nodes $(D_1, ., D_n)$. Note that, network coding is only applied to the flows passed through the paths correspond to the target permutation and the packets of other flows are processed via regular routing. For example, (3,1,2) is the target permutation of the scheme in Fig. 1a. Note that, for the n-input n-output star structure, there are $n!$ distinct states for the target permutation. However, each of the $n!$ target permutations does not necessarily lead to a Star-NC scheme of size n. Indeed, for states that direct links between some of input and output nodes exist, the coding scheme converts to a partial Star-NC of smaller size plus a set of direct transmissions between input and output nodes. From a practical point of view, this scenario is not happened for unicast sessions which are routed by an algorithm with shortest-path metric. For example, routing of three unicast sessions for the target permutations (2,3,1), (3,2,1) and (2,3,1) corresponds to a full-star NC of size 3, whereas for (2,1,3) and (1,3,2), it leads to a partial-star NC of size 2 plus a direct link between an input and output node. Furthermore, the permutation (1,2,3) is a trivial case since the Star-NC has no benefits over traditional routing done by seven unicast transmissions.

Further, each node is constrained to be in the transmission range of the nodes that overhears them. We assume that for a full-star of size 3, each node is capable of overhearing of a neighbor node in both clockwise and counter clockwise directions, i.e. a wheel topology. In general, this assumption for a full-star of size n holds for overhearing of m neighbors in both directions. We referred to m as the *radius of overhearing*. It is simple to verify that $m \geq n/3$ for a star structure of size $n$with a regular pattern, i.e. each node is within the communication range of at least $n/3$ of neighbors in both directions.

It is not necessary for the Star-NC that the set of input/output nodes have distinct elements, i.e. some of nodes are the same. This means that the paths, correspond to multiple unicast sessions, are not necessarily *edge-disjoint*, i.e. there is an edge which lies in more than one path. Thus a structure of size n could have less than $2n$ distinct nodes. For example, in the Star-NC of Fig. 1a, locating D_2, in the same position as D_1, results a star structure of size 3 with 5 nodes. The same thing could happen if S_1 and S_2 have the same position where is located in the transmission range of both D_1 and S_3. Thus, taking a specific Star-NC scheme as basis, the overlapping of the nodes allows the creation of various coding schemes of smaller size.

2.2 Uniqueness of Star-NC

Wireless networks exhibit significant data redundancy [19], i.e., there is a large overlap in the information available to the nodes. First, as a packet travels multiple hops, its content

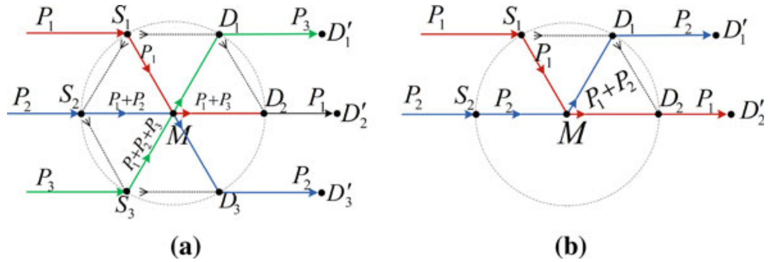

Fig. 1 The Star-NC scheme. **a** A full-star structure of size 3. **b** A partial-star structure of size 2

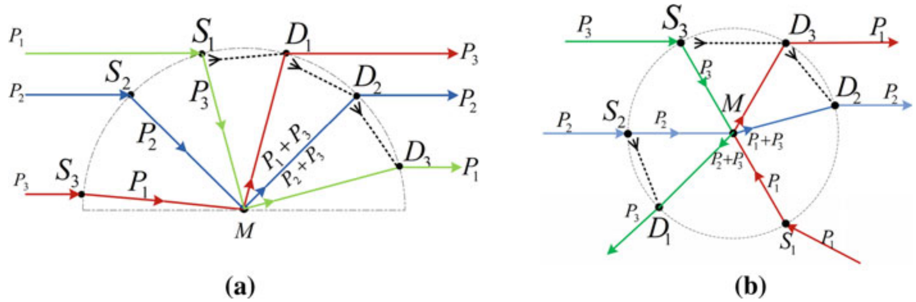

Fig. 2 Coding scheme proposed by Theorem 1. **a** A partial-star structure of size 3. **b** An extended form of partial-star structure of size 3

becomes known to many nodes. Furthermore, wireless broadcast amplifies this redundancy since, in each hop, it delivers the same packet to multiple nodes within the transmitter's radio range. Network coding can take the advantage of data redundancy along with broadcast nature of wireless medium to reduce the number of transmissions. Thus the more flexible ways of overhearing a coding has, the more opportunities of coding will be created.

Our scheme has unique features compared to COPE-type coding schemes. From opportunistic listening view, our scheme takes advantage of the proximity of all the nodes around the relay node, i.e. the combinations (next-hop, previous-hop), (next-hop, next-hop) and (previous-hop, previous-hop) are legal for listener and overheard nodes. However, the opportunistic listening for COPE-type NC is only limited to the case (next-hop, previous-hop) [7].

Precisely in addition to (next-hop, previous-hop), *two-hop* Star-NC has the case (next-hop, next-hop) and *three-hop* Star-NC has both the combinations (previous-hop, previous-hop) and (next-hop, next-hop) for opportunistic listening. These extra forms of opportunistic listening allow the nodes to use the broadcast nature of wireless medium more efficient than COPE.

For example, the simple scheme in Fig. 1b, benefits from the closeness of D_1 and D_2 with opportunistic listening. This feature identifies the Star-NC scheme from a COPE-type coding scheme. A more sophisticated structure is shown in Fig. 2b where the geographical positions of the nodes limits the overhearing only to D_1 from S_1, D_3 from S_3 and D_2 from D_3. Note that the overhearing D_2 from D_3 has a critical role in coding opportunity, i.e., if the overhearing, like COPE, is limited to the case output-from-input, no coding is applicable for this structure. On the other hand, employing the (previous-hop, previous-hop) overhearing and sending coded packet, instead of native packet, to the relay node makes more data redundancy, e.g. in Fig. 1a, opportunistic listening of S_2/S_3 from S_1/S_2 brings the information of P_1 and P_2 to D_3

2.3 Practical Issues

It is worth mentioning that two-hop Star-NC, such as Fig. 1b, does not require any coordination among nodes since no coding is performed at input nodes. Therefore, like COPE, node M is the only coder node and thus can make decision of NC with the received packets from input nodes. In summary, the NC opportunities for each node such as M are identified by its local information.

On the other hand, in three-hop Star-NC (such as Fig. 1a), we need a type of coordination among the input nodes and the relay node. To follow the first step of coding scheme, the input nodes must be able to detect the flows belonging to specific unicast sessions. It is clear that the implementation of flow coordination in a general wireless network is a more challenging task. Thus if the cost of flow coordination is not acceptable, we can limit the use of Star-NC to two-hop coding schemes. Next we will show that this group (two-hop schemes) is more practical in WMN and itself contains a large portion of Star-NC opportunities. However, we think that the flow coordination is practical for specific networks such as wireless mesh network. Indeed a wireless mesh network has the following features:

1. The topology is relatively stable except for the occasional failure of nodes or addition of new ones.
2. Traffic flow, aggregated from a large number of end users, changes occasionally, and hence could be assumed constant for an extended period of time, i.e. the set of unicast pairs remain stable over a large time-scale.

Thus an inter-flow coordination among a set of nodes is applicable for a large period of time. This allows for the nodes to re-coordinate when the flow structure changes. Flow coordination could be triggered by the relay node when finds a suitable Star-NC scheme. After that, to be able to send its packet (likely coded) to the relay node, each input node must be capable of detecting the packet of flows belonged to Star-NC scheme. In the case of source routing protocol, it is simply accessible by header information of each packet. Otherwise, it could be provided by means of the relay node's route table, i.e. the relay node sends its route table to the input nodes when it triggers a specific Star-NC scheme.

The above explanations may convince us to accept the cost of flow coordination in wireless mesh network. Anyhow, for understanding the tradeoff between Star-NC gain and coordination drawback, we consider two distinct scenarios for Star-NC relative to two/three-hop coding schemes in our evaluations.

3 Star-NC Analysis

The throughput improvement of Star-NC depends on the existence of coding opportunities, which themselves depend on the network topology and traffic patterns. This section provides a deep insight into the expected throughput improvement and the factors affecting it for a network consists of a single star topology. At the next section, we study the coding effect on aggregated throughput of a general network by means of an LP formulation. In other word, through this section we analyze the coding gain for a single star structure while at the next section we study the throughput performance of coding for a general wireless network consists of multiple star structures.

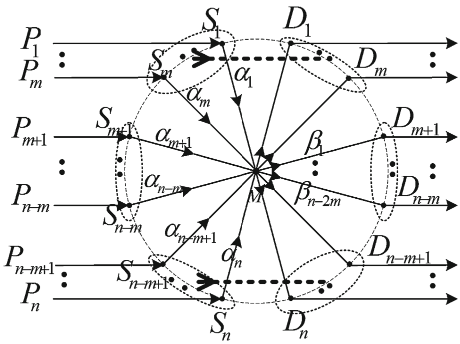

Fig. 3 Lower bound of $N(n, m)$, at least $n - 2m$ transmissions are required to be transmitted to output nodes via M

3.1 Definition and Modeling Assumptions

Throughout the paper, we use n and m, respectively, as the size of full-star structure and the radius of overhearing. The size of a partial-star structure is denoted by η. We model the star structure of size n with a star topology whose degree at center node is equal to $2n$, i.e. $2n$ nodes are around the relay node. The nodes are located in order that (1) a line passing through the center can partition the input and output nodes and (2) each node is capable of overhearing m neighbor nodes. As mentioned earlier, overhearing for partial star is limited to one of clockwise or counterclockwise direction while for full star can be done in both directions. A typical full-star structure of size n is shown in Fig. 3.

The rest notations are defined in Table 1. Briefly, we formulize the Star-NC problem in two encoding matrices, namely $\mathbf{Q_B}$ and $\mathbf{Q_M}$, which, respectively identifies the coding at input nodes and relay node. Thus the main problem can be represented as finding the encoding matrices. For example, the encoding matrices for the star structures in Fig. 1a, b, are:

	Figure 1a	Figure 1b
$\boldsymbol{Q_B}$	$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$
$\boldsymbol{Q_M}$	$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \end{pmatrix}$

Due to the simplicity of overhearing in partial star, we first analyze the bounds of coding gain for this structure and then extend the analysis to the full-star structure. In a full-star structure, according to Fig. 3, two groups of output nodes can overhear the input nodes: (1) output nodes at the top of the star (i.e., D_1 through D_m) and (2) output nodes at the bottom of the star (i.e., D_{n-m+1} through D_n). In contrast, in the partial-star structure whose set of input and output nodes covers only a portion of the circle around the relay node, opportunistic listening of the input nodes only could be done by the first group of the output nodes (i.e., D_1 through D_m).

Next, we provide a definition for the coding gain and then analyze the Star-NC gain for partial and full star-structures, respectively in Sects. 3.2 and 3.3.

Table 1 Notations and variables used in coding analysis

$G(n, m)$	The gain of the Star-NC scheme
$N(n, m)$	The number of required (coded/uncoded) packets to be sent from the relay node to output nodes
$N_{Min}(n, m)$	The minimum values for $N(n, m)$
$N_{Max}(n, m)$	The maximum values for $N(n, m)$
$\pi(a_1, a_2, \ldots, a_n)$	A specific permutation of $(a_1, a_2, \ldots, a_n)$. Further, π and π_n without any arguments denote any permutation of $(1, 2, \ldots, n)$
$N_\pi(n, m)$	The value of $N(n, m)$ for a specific target permutation π
P_i	The incoming packet of input node S_i. Further, the XOR of P_i and P_j is denoted by $P_i + P_j$
$\boldsymbol{P} = [P_1 P_2 \ldots P_n]^{\boldsymbol{T}}$	Vector of the incoming packet to the star structure
α_i	A non-empty XORed subset of $\{P_1, P_2, \ldots, P_n\}$ sent by an input node S_i to the relay node $(1 \le i \le n)$
β_j	A non-empty XORed subset of $\{P_1, P_2, \ldots, P_n\}$ sent by the relay node to output nodes $(1 \le j \le N_\pi(n, m))$
$\mathbf{Q}_{\boldsymbol{B}}(\boldsymbol{n} \times \boldsymbol{n})$	Encoding matrix of input nodes $(\mathbf{Q}_{\boldsymbol{B}} \times \boldsymbol{P} = [\alpha_1 \ldots \alpha_n]^{\boldsymbol{T}})$
$\mathbf{Q}_{\boldsymbol{M}}(N_\pi(n, m) \times \boldsymbol{n})$	Encoding matrix of the relay node for target permutation π $(\mathbf{Q}_{\boldsymbol{M}} \times \boldsymbol{P} = [\beta_1 \ldots \beta_{N_\pi(n,m)}]^T)$
$\mathbf{I}_{\boldsymbol{n}}$	The Identity matrix of size n. $\mathbf{Q}_{\boldsymbol{B}} = \mathbf{I}_{\boldsymbol{n}}$ means no encoding at the input nodes
V	The set of nodes in the network
E	The set of directed links in the network

3.1.1 Coding Gain

As mentioned above, the problem of Star-NC can be represented as finding the optimum encoding matrices. The optimality of coding is identified by definition of the coding gain. The effectiveness of a Star-NC scheme can be explained either by the number of required transmissions and the number of opportunistic listening. Note that each target permutation in Star-NC has different coding in terms of the number of packet transmissions, the number of coded packets and the overhearing pattern. It is an interesting problem to find the most effective network coding for each target permutation.

We define the *coding gain* as the ratio of the number of transmissions required by the current non-coding approach, to the number of transmissions used by Star-NC to deliver the same set of packets from the input to the output of the structure. For example, the gains of the schemes in Fig. 1a, b are equal to $\frac{9}{7}$ and $\frac{6}{5}$, respectively. In COPE, there are two methods of gain evaluation which are referred as *Coding Gain* and *Coding+MAC Gain* [7]. The *Coding+MAC Gain* only concentrates on the relay node of the structure and defined as the rate improvement of the relay node in the output links. Our gain definition corresponds to *Coding Gain* method which considers all the nodes of the structures. Note that this definition is local metric that considers only a set of specific flows as well as the nodes in the structure. Thus, the relation of this metric and a global objective such as end-to-end throughput is not straightforward. This metric is equal to throughput improvement if we assume a network with a single star structure and only with its corresponding unicast flows. Further, it is supposed that the all of incoming flows have the same rate. At the next section, we provide an LP formulation to find the effect of the coding on aggregated throughput of a general network.

3.2 Partial Star-NC

The gain for partial-star structure is only dependent on the value of η and m, i.e. is constant for different target permutations. We begin with the following theorem which identifies an upper bound for the number of relay node's transmissions.

Theorem 1 *For a partial-star structure of size η and a radius of overhearing m nodes, in which $m \leq 4$ and $m \leq \eta \leq 3m+2$, the value of $N(\eta, m)$ is upper bounded by:*

$$N(\eta, m) \leq \eta - m \tag{1}$$

Proof The proof, detailed in Appendix A, uses a fixed $\mathbf{Q}_{\boldsymbol{B}}$ which equals to $\mathbf{I}_{\boldsymbol{n}}$ for $m = 1$. For $m > 1$, we use a fixed $\mathbf{Q}_{\boldsymbol{B}}$ as an upper triangular matrix whose non-zero elements are in the main diagonal and $km + 1$th rows for which $0 \leq k < [n/m]$. Furthermore, the core of $\mathbf{Q}_{\boldsymbol{M}}$ is consisted of an identity matrix of size $\eta - m$ concatenated by a matrix $J_{(\eta-m)\times m}$ whose columns have even number of ones. Based on the target permutation, $\mathbf{Q}_{\boldsymbol{M}}$ with exchanged columns must be used. Thus for $m > 1$, $\mathbf{Q}_{\boldsymbol{M}}$ depends on the target permutation while for $m = 1$ is a fixed $(\eta - 1) \times \eta$ matrix. This leads to a constructive proof which proposes a proper network coding scheme that achieves the desired gain.

An example of the coding proposed by Theorem 1 is shown in Fig. 2a for $\eta = 3, m = 1$ and target permutation (3,2,1). Now, we are ready to study the tightness of this upper bound, i.e., we evaluate a lower bound for $N(\eta, m)$ and accordingly the coding gain of partial Star-NC.

Preposition 1 *The coding gain of a partial Star-NC of size η with $m \leq 4$ and $m \leq \eta \leq 3m+2$ is equal to $\frac{\eta}{\eta-m}$.*

Proof With the assumption of fixed m, it is simple to verify that $N(\eta, m) \geq \eta - m$, too.Since each output node catches exactly m packets by overhearing, in general at least $\eta - m$ packets (likely coded) are required for each output nodes to be able to solve a system of η linear equations. These extra packets are only provided by the relay node. Thus by considering both inequalities we can see that $N(\eta, m) = \eta - m$ and thus we will have $G(\eta, m) = \frac{\eta}{\eta-m}$. □

Obviously this gain is achievable for any target permutation. Further as $N(\eta, m) \geq \eta - m$, the coding proposed by Theorem 1 has the maximum gain. The assumption $m \leq 4$ in Theorem 1 is required due to the largeness of problem state in which makes our proof is unfeasible for $m > 4$. Although we have not proved, we have sufficient evidence to conjecture that the previous theorem is also held for other values of m. That is, we have the following statement.

Conjecture 1 *Theorem 1 holds for all m for which $m \leq \eta \leq 3m + 2$.*

Notwithstanding we assume that each node is capable of overhearing m neighbors, the coding of Theorem 1 does not use all opportunities of overhearing, e.g. in Fig. 2a, S_1 and S_2 do not overhear any node. A more controversial assumption for the theorem is that the radius of overhearing is not fixed to m, i.e. it is equal to m for some of the nodes and is greater than m for the others. We did not discuss such cases.

3.2.1 Partial Star Structure with $m = 1$

An interesting result of Theorem 1 is that the partial Star-NC with $m = 1$ is a *two-hop* coding scheme since $\mathbf{Q_B} = \mathbf{I_n}$ for $m = 1$, i.e., no coding at input nodes. We engage in two-hop Star-NC schemes since they do not arise any flow coordination for the nodes. Furthermore, the proposed coding does not imply any condition on the position of input nodes, i.e. no input node is constrained to be in the communication range of the other input nodes. These lesser conditions make the partial Star-NC with m = 1 to be more practical in WMN, i.e. the opportunities for partial Star-NC is more than full Star-NC in a real network.

Furthermore, $\mathbf{Q_M}$ is a constant matrix for each target permutation, i.e. the relay node sends $P_1 + P_n, P_2 + P_n, \ldots, P_{n-1} + P_n$ to output nodes. Thus each output nodes require catching a single packet via overhearing to be able to decode its packet. The overhearing can be done from either an input or output node. This condition can be satisfied with a more general form of star structure, i.e., it is not necessary for the nodes to partition into input/output groups with a line passing through the relay node. For example, this coding is applicable to the structure shown in Fig. 2b where the geographical positions of the nodes limits the overhearing only to D_1 from S_1,D_3 from S_3 and D_2 from D_3. Note that the overhearing D_2 from D_3 has an important role in coding, i.e., if the overhearing like COPE is limited from the output nodes to the input nodes, no coding is applicable for this structure.

3.3 Full Star-NC Analysis

Now, we are ready to study the gain of network coding for a full-star structure. As we will see, the coding gain for full star structures (unlike the partial ones) strictly is dependent on the target permutation and thus is not fixed. In this section, we try to find both the bounds of the gain and the permutations with maximum coding gain. Next, we study the coding for a general target permutation and expected value of the gain. Lastly, we propose an algorithm that finds Star-NC for each target permutation.

3.3.1 Bounds of the Gain

Theorem 2 *The value of $N(n, m)$ for a full-star structure of size $n \geq 3$ with $\lfloor n/3 \rfloor \leq m < \lfloor n/2 \rfloor$ has the following bounds:*

$$n - 2m \leq N(n, m) \leq n - m \tag{2}$$

Proof By considering the radius of overhearing, at most $2m$ of the input nodes could be opportunistically listened by some of the output nodes. This is shown in Fig. 3 where some of the nodes S_1 through S_m (and similarly S_{n-m+1} through S_n) are overheard by $D_1, \ldots, D_m$ (and $D_{n-m+1}, \ldots, D_n$), respectively. Thus, even if we assume that the output nodes know the information of each other, at least $n - 2m$ packets must yet be received by the output nodes via M to solve a system of n liner equations. Alternatively, if we focus on output nodes D_{m+1} through D_{n-m}, we can see that these nodes do not overhear any packet of input nodes and can only overhear other output nodes which decode and forward their packets. Thus, at least $n - 2m$ encoded packets such as $(\beta_1, \ldots, \beta_{n-2m})$ are required by D_{m+1}through D_{n-m} for decoding their packets. On the other hand, the upper bound of (2) is accessible by means of Theorem 1. It is sufficient to consider n input and n output nodes on an arc with radius of overhearing m nodes. Clearly, Theorem 1 provides a coding scheme for this structure with clockwise overhearing. □

Corollary 1 *The gain of the Star-NC scheme for a full-star structure has the following bounds:*

$$\frac{3n}{3n-m} \leq G(n,m) \leq \frac{3n}{3n-2m} \tag{3}$$

Proof Using the definition of Star-NC gain, we have:

$$G(n,m) = \frac{3n}{2n+N(n,m)} \tag{4}$$

Then, by applying the bounds for $N(n,m)$ in (2), the result (3) is simply obtained. □

The case $n = 3$ is a special type of the star structure which leads to a fixed maximum achievable gain for all non-trivial target permutations. As mentioned above, out of the six different target permutations, the coding scheme for three corresponds to a full-star of size 3, for two leads to a partial-star of size 2 plus a direct transmission, and for one of them has no benefits over traditional routing. As this gain is equal to the maximum achievable gain in (3), we discuss it separately in the following preposition.

Preposition 2 *The gain of Star-NC scheme for a full-star structure of size 3 with $m = 1$ is equal to $\frac{9}{7}$.*

Proof It is enough to show that $N_\pi(3,1) = 1$. For π which leads to a full-star structure, we propose the following schemes:

π	(3,1,2)	(3,2,1)	(2,3,1)
Q_B	$\begin{pmatrix} 1&0&0\\1&1&0\\1&1&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1\\0&1&0\\0&1&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1\\0&1&1\\0&0&1 \end{pmatrix}$
Q_M	$\begin{pmatrix} 1&0&1 \end{pmatrix}$	$\begin{pmatrix} 1&1&1 \end{pmatrix}$	$\begin{pmatrix} 1&0&1 \end{pmatrix}$

As the gain for full-star NC is equal to the right hand of (3), this is the maximum achievable gain and thus the proposed coding schemes are optimal Star-NC schemes. □

3.3.2 Permutation with Maximum Star-NC Gain

We are interested in finding scenarios in which the maximum and minimum Star-NC gain is achieved. Furthermore, we need to identify other target permutations which lead to a specific gain other than the minimum and maximum identified by (3). We propose to find a good upper bound for the gain of such permutations by examining the differences between these permutations and those that provide the maximum gain.

Definition 1 (*Parallel permutation*): A parallel permutation $\pi(1,2,\ldots,n)$ is a permutation which can be partitioned into two consecutive permutations of $(1,2,\ldots,k)$ and $(k+1,\ldots,n)$, respectively, where: $\lfloor n/3 \rfloor \leq k \leq n - \lfloor n/3 \rfloor$.

Definition 2 (*k-Interleaved permutation*): a permutation is k-interleaved when $n = k\acute{n}$ and $m = k\acute{m}$ and for each $1 \leq i \leq k$ and $1 \leq j \leq \acute{n} - 1$ its elements in positions $i + jk$ are congruent to $i - 1$ modulo k ($m = \lfloor n/3 \rfloor$, $\acute{m} = \lfloor \acute{n}/3 \rfloor$). In this case for each i, we name the elements placed in the positions $i + jk$ as the ith sub permutation and denote it as $\pi^{(i,k)}$.

An interesting features of k-interleaved permutation is that the corresponding Star-NC scheme of size n can be divided into k distinct schemes of size n/k. Indeed, the above condition implies that, for each i, the source and destination nodes of $\pi^{(i,k)}$have a fixed distance from each other and thus makes a regular star structure of size n/k. Now, we are ready to introduce the permutations which achieve the maximum gain of network coding in star structures. We will refer to these permutations as *max-gain* permutations.

Lemma 1 (Max gain lemma)*: the maximum Star-NC gain is achieved for a target permutation which is: I)a parallel permutation, II)a k-interleaved permutation whose sub permutations have size 3 and III)A k-interleaved permutation whose sub permutations belong to any of I and II.*

Proof For (I), we can divide the problem of full-star NC of size n into two disjoint coding for partial-star NC schemes of size k and $n-k$, respectively. Opportunistic listening for the former ($\pi_1(1,\ldots,k)$) is done in the clockwise direction while for the latter ($\pi_2(k+1,\ldots,n)$) is performed in the counter clockwise direction. Assume m is the radius of overhearing, Theorem 1 implies that network coding for each partial-star saves m transmissions of the relay node. Thus, for $N_\pi(n,m)$, we have:

$$\begin{aligned} N_\pi(n,m) &= N_{\pi_1}(k,m) + N_{\pi_2}(n-k,m) \leq (k-m) + (n-k-m) \\ &= n-2m \end{aligned}$$

Further, the lower bound of Theorem 1 implies that $N_\pi(n,m) \geq n-2m$. Hence, $N_\pi(n,m) = n-2m$

For the case of k-interleaved permutation, we can divide the problem of the Star-NC into k different schemes of sub permutations $\pi^{(1,k)}$ through $\pi^{(k,k)}$. For both cases (II) and (III), each sub permutation corresponds to a Star-NC scheme with the maximum gain. Thus, by the mathematical induction, we can obtain:

$$\begin{aligned} N_\pi\ (n,m) &= \sum_{i=1}^{k} N_{\pi^{(i,k)}}(\acute{n},\acute{m}) = k\left(\acute{n}-2\acute{m}\right) = k\acute{n}-2k\acute{m} \\ &= n-2m \end{aligned} \tag{5}$$

This completes the proof. □

For example, this lemma states that both the permutations (1,3,2,5,4,6) and (6,5,4,3,2,1) are max-gain permutations. The former is a parallel permutation, while the later is a 2-interleaved permutation whose Star-NC scheme could be converted into two distinct Star-NC schemes of size 3. By applying Preposition 2 twice, we can find an optimum network coding scheme.

Further, an important result of this lemma is that for case (I), the resulting scheme will be a two-hop Star-NC if the corresponding partial-star schemes are two-hop Star-NC. The same is true for case (III) where its sub permutations lead to two-hop Star-NC schemes.

3.3.3 The Gain for General Target Permutations

We are interested in computing the gain of a non max-gain permutation. Therefore, we try to define a metric which provides the difference between a general permutation and a permutation with the maximum gain.

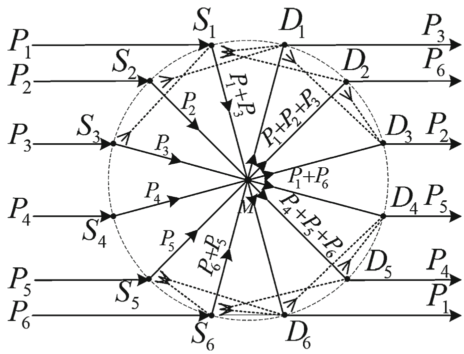

Fig. 4 Star-NC scheme of $\pi = (3, 6, 2, 5, 4, 1)$ by reusing the scheme of $\sigma = (3, 1, 2, 5, 4, 6)$and an additional transmission as $P_1 + P_6 (m = 2)$

Definition 3 (*Permutation distance*) The distance between two permutations P_A and P_B, denoted by $\mathcal{D}(P_A, P_B)$, is the minimum number of pair exchanges which converts P_A to P_B. The gain distance of P_A, denoted by $\mathcal{D}(P_A)$, is the minimum distance of P_A and a max-gain permutation.

For instance, the distance between $P_A = (1, 5, 3, 4, 6, 2)$ and $P_B = (1, 2, 3, 4, 6, 5)$ is equal to one since we can reach from P_A to P_B by exchanging P_2 and P_5. As P_B is a parallel permutation, $\mathcal{D}(P_A) = 1$.

Lemma 2 (Distance lemma)*: The value of $N\ (n, m)$ for a general permutation π is upper-bounded by:*

$$N_\pi\ (n, m) \leq N_{Min}\ (n, m) + \mathcal{D}\ (\pi) \tag{6}$$

Proof Let $k = \mathcal{D}(\pi)$, Assume k pairs of exchanges (i_1, j_1),(i_2, j_2),...,(i_k, j_k) convert π to a max-gain permutation. We set M to send k encoded packets $P_{i_1} + P_{j_1}$,$P_{i_2} + P_{j_2}$,...,$P_{i_k} + P_{j_k}$ to the output nodes, in addition to those it sends for the corresponding max-gain permutation. With this new scheme, the packet decoding for π at the output nodes is similar to its corresponding max-gain permutation except the states that P_{j_q} is available whereas we need to know P_{i_q} $(1\leq q \leq k)$. In these situations, we recover P_{i_q} by XORing the additional encoded packet $(P_{i_q} + P_{j_q})$ with P_{j_q}. □

An interesting result of this lemma is that if the scheme of σ corresponds to a two-hop Star-NC, the resulting scheme for π will be a two-hop Star-NC, too.

As an example, let $\pi = (3, 6, 2, 5, 4, 1)$. Then $\mathcal{D}(\pi) = 1$ since $\mathcal{D}(\pi) = \mathcal{D}(\pi, \sigma)$ for $\sigma = (3, 1, 2, 5, 4, 6)$ as a parallel permutation. This lemma proposes a scheme (Fig. 4) for π which is consisted of the scheme for σ plus the encoded transmission $P_1 + P_6$. This additional coded packet is formed by using the distance of π and σ. Note that, in the decoding step of π, this packet is used by D_2 and D_6 to recover P_6 and P_1, respectively.

Minimum Gain: Lemma 1 offers the minimum gain to the permutation that has longest distance from a max-gain permutation. We refer to these permutations as m*in-gain permutations*. Obviously, the maximum distance is equal to $\lfloor m = n/3 \rfloor$ since each permutation can be converted to a parallel one by at most m exchanges. The permutations (3,4,5,6,2,1) and (5,6,1,3,2,4) are examples of min-gain permutations.

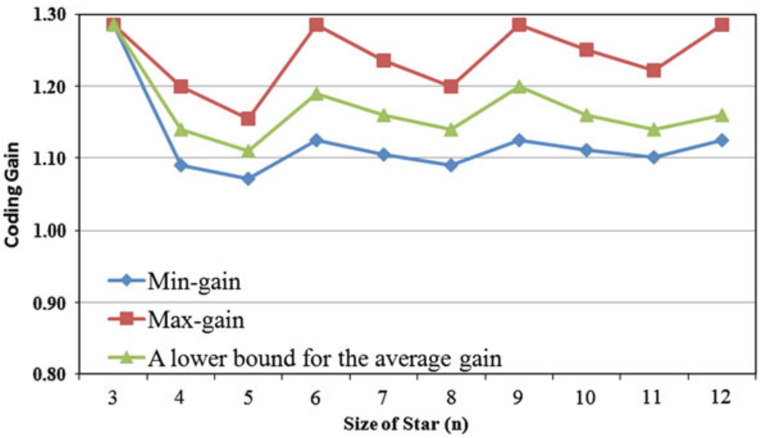

Fig. 5 Bounds of the Star-NC gain along with a lower bound on the expected value of the gain (full-star structure)

3.3.4 A Lower Bound on the Expected Value of the Gain

Let $S_i(n)$ denotes the number of permutations with a gain distance equals to i. It implies that $S_0(n)$ and $S_m(n)$ equal to the number of max-gain and min-gain permutations, respectively. Now, we can calculate an upper bound for the expected value of $N(n, m)$ based on the value of $S_i(n)$ for $0 \leq i \leq m$. We classify all $n!$ permutation into $(m+1)$ groups based on the gain distance as:

$$E[N(n,m)] = \sum_{i=0}^{m} P(\mathcal{D}(\pi) = i) \times N_{\pi}(n,m) \leq \sum_{i=0}^{m} P(\mathcal{D}(\pi) = i) \times (N_{Min}(n,m) + \mathcal{D}(\pi)) \tag{7}$$

Since π denotes a random selected permutation, we have:

$$E[N(n,m)] \leq \sum_{i=0}^{m} \frac{S_i(n)}{n!} \times (N_{Min}(n,m) + i) \tag{8}$$

Thus, we obtain a lower bound for the expected value of $G(n, m)$ as:

$$E[G(n,m)] = E\left(\frac{3n}{2n + N(n,m)}\right) \geq \sum_{i=0}^{m} \frac{S_i(n)}{n!} \times \frac{3n}{2\text{n} + N_{Min}(n,m) + i} \tag{9}$$

Finding the values of $S_i(n)$ is a sophisticated process. We found an explicit formula for $i = 0$ while the other values are calculated by a computerized program. The corresponding gains for min and max-gain permutations together with the lower bound of $E[G(n, n/3)]$ are depicted in Fig. 5.

3.3.5 Finding Star-NC Scheme for Each Permutation

By the means of Lemmas 1 and 2, we can find the gain of Star-NC for a desired permutation. Now, we are equipped to propose an algorithm that finds the Star-NC scheme for each target permutation. Algorithm 1 represents the pseudo code of finding the coding matrices, $\mathbf{Q}_M$

and $\mathbf{Q}_B$, for the specific permutation π_n. The algorithm offers a two-hop Star-NC ($\mathbf{Q}_B = \mathbf{I}_n$) for states 2, 3 and 4 where the corresponding sub-schemes are two-hop Star-NC. More specifically, the situations are occurred either both π_1 and π_2 (in state 2) or each $\pi_n^{(i,k)}$(in state 3) or σ (in state 4) correspond(s) to a two-hop Star-NC. Furthermore, since the partial Star-NC with $m = 1$ is always a two-hop coding scheme, we can state that the Star-NC for each π_n with n>3 and $m = 1$ is a two-hop coding scheme.

Inputs : π_n
Outputs : Coding matrices (Q_M, Q_B) for π_n
Find-Star-NC-Scheme(π_n: permutation)
1. if $n = 3$ return (Q_M, Q_B) as in Preposition 2;
2. if π_n is a parallel permutation as $\pi_1(1,..,k)\pi_2(k+1,..,n)$
   Find (Q_{M_1}, Q_{B_1}) of π_1 using Theorem.1
   Find (Q_{M_1}, Q_{B_1}) of π_2 using Theorem.1
   Set $Q_M = \begin{pmatrix} Q_{M_1} & 0 \\ 0 & Q_{M_2} \end{pmatrix}$ $Q_B = \begin{pmatrix} Q_{B_1} & 0 \\ 0 & Q_{B_2} \end{pmatrix}$
3. else if π_n is a k-interleaved permutation of size 3 or parallel sub-permutations
   Find $(Q_M^{(i,k)}, Q_B^{(i,k)})$ for $\pi_n^{(i,k)}$ for $1 \leq i \leq n/k$ using mentioned methods 1 and 3;
   Set Q_B as interleaved result of k matrices $Q_B^{(i,k)}$
   Set Q_M as interleaved result of k matrices $Q_M^{(i,k)}$
4. else
   Find max-gain permutation σ which $k = \mathcal{D}(\pi_n, \sigma)$
   Assume k exchanges (i_1, j_1), (i_2, j_2),…,(i_k, j_k) convert $\pi_n$to σ
   Find (Q_M^σ, Q_B^σ) of σ using mentioned methods 2 and 3;
   Set $Q_B = Q_B^\sigma$
   Generate $Q_M$from Q_M^σ by adding k rows related to encoded packets:
   $P_{i_1} + P_{j_1}$, $P_{i_2} + P_{j_2}$,…, $P_{i_k} + P_{j_k}$
   end if;
return(Q_M, Q_B);

4 Star-NC Optimization Framework for General Network Topologies

In this section, we formulate a linear programming (LP) framework to find the maximum throughput of the network using Star-NC scheme. The framework uses an LP technique similar to ones used in [8,12,13]. But our framework significantly differs from [8] which has a two-hop coding scheme with opportunistic listening while Star-NC has both two and three hop coding schemes. Further, unlike the systems in [12] and [13] which consider only network coding in the absence of opportunistic listening; our scheme uses opportunistic listening effectively.

The notations and modeling assumptions are listed in Table 2. We use the protocol model of interference introduced by Gupta and Kumar [15], i.e., two nodes have a link if their distance is less than *communication range* and are interfered if their distance is less than *interference range*. Also, two links $e_1 = (i_1, j_1)$ and $e_2 = (i_2, j_2)$ are interfered either as j_1 is within the interference range of i_2 or j_2 is within the interference range of i_1.

4.1 Star-NC Modeling

We present Star-NC by a five tuple$(M, \mathcal{S}, \mathcal{D}, \pi, m)$ and denote it as ξ, where M is the relay node, π is the target permutation and m is the radius of overhearing. Further $\mathcal{S}$ and $\mathcal{D}$,

Table 2 Variables and notations used in system modeling

$E^{+}(v)$	The set of incoming links incident on node v
$E^{-}(v)$	The set of outgoing links incident on node v
D	The set of traffic demands (related to unicast sessions)
$D(k)$	Traffic amount which requested by session k
λ_k	The end-to-end throughput for the demand k, i.e. a portion of $D(k)$ which can be routed by the network
$C(e)$	Capacity of data link e
$t(e)$	The transmitting node of the directed link e
$r(e)$	The receiving node of the directed link e
$s(k)$	Source node of traffic demand k
$d(k)$	Destination node of traffic demand k
$\mathcal{S}$	Data links of star structure from input nodes to relay node, i.e. $\mathcal{S} = \{S_1, \ldots, S_n\} \times \{M\}$
$\mathcal{D}$	Data links of star structure from relay node to output nodes, i.e. $\mathcal{D} = \{M\} \times \{D_1, \ldots, D_n\}$
$\xi = (M, \mathcal{S}, \mathcal{D}, \pi, m)$	Star structure of size n by five tuples: M is the relay node, $\mathcal{S}$ and $\mathcal{D}$, respectively, denote the input and output links to M, π is target permutation for routing and m is radius of overhearing $(n = \|\mathcal{S}\| = \|\mathcal{D}\|)$.
Γ	Set of all star structures which employ Star-NC scheme
$f\ (e)$	Total flow rate passed through link e
$F_k(P)$	The flow rate of demand k over path P
P_k	The set of available paths for demand k
$f^{NC}\ (\xi)$	Flow rate of coded traffic for star ξ which broadcast by M to output nodes
$z_e^k\ (P)$	The portion of the traffic on path P for demand k that is transmitted as uncoded from link e .
$E^C\ (e)$	The set links which conflict by link e
$\Gamma^C(e)$	The set Star structures which conflict by link e

respectively, represent the incoming and outgoing links of M, i.e. $\mathcal{S} = \{S_1, \ldots, S_n\} \times \{M\}$ and $\mathcal{D} = \{M\} \times \{D_1, \ldots, D_n\}$ where n denotes the size of star ($n = |\mathcal{S}| = |\mathcal{D}|$). We refer to the component x of a specific ξ as $x(\xi)$, e.g. $M(\xi)$,$\mathcal{S}(\xi)$,$\mathcal{D}(\xi)$ and $n(\xi)$. Further, $N(\xi)$ denotes $N_\pi\ (n, m)$. In this modeling the following conditions must be satisfied:

1. Set of n unicast transmissions intersect each other at M and are routed to output nodes according to target permutation π.
2. Set of input/output nodes must be located on the star in order that each node is within the transmission range of the nodes that overhears them.

Above modeling allows Star-NC schemes to overlap with each others. We use a straightforward method to generate all the valid Star-NC structures for a given network. Let d_i denotes degree of node i. The number of different stars of size n with i as the relay node (ignoring the target permutation and the classification of neighbors into input/output nodes) is upper bounded by $\binom{d_i}{2n}$. Thus for a given network $G(V, E)$, the number of star structures of size n is bounded by:

$$\sum_{i \in V} \binom{d_i}{2n} \leq \sum_{i \in V} \frac{d_i^{2n}}{(2n)!} \leq \frac{1}{(2n)!} \left(\sum_{i \in V} d_i \right)^{2n} = \frac{4^n}{(2n)!} |E|^{2n} \tag{10}$$

Hence, the number of star structure is $O\left(|E|^{2n}\right)$. Since nodes in a wireless mesh network have small degrees, the size of star (n) would become relatively small. Therefore, it is relatively simple and fast to generate all such star structures, which we refer to them as *star*-bases. For each star-basis, we can generate all possible Star-NC combinations by labeling its nodes as input/output nodes. Let Γ denote the set of all valid Star-NC opportunities. Note that the conditions (1) and (2) are only depend on the network topology, i.e. independent of the traffic model and the routing strategy. Thus for a given network, we can compute Γ only once in the pre-computation stage. After identifying the traffic and routing, we can check the condition (3) for each element of Γ and include it in our LP formulation if the condition is satisfied.

4.2 Problem Formulation

We model the traffic as a set of traffic demand denoted by D. The demand k corresponds to $D(k)$ amount of traffic (e.g. in Mbps) requested by source node $s(k)$ to be routed to destination node $d(k)$. As in [8,12,13], we define the throughput as a multiplier λ such that for each demand k, at least $\lambda D(k)$ amount of requested traffic is guaranteed to be routed by the network. This definition holds the linearity of the system while provides a means of fairness. Note that the aggregated network throughput is equal to the sum of all routed traffic for each demand, i.e. $\sum_{k \in D} D(k)\lambda$. We have the following constraints.

Fairness constraint: In our system, we consider multi-path routing. Let P_k be the set of available paths for the routing demand k from $s(k)$ to $d(k)$. Assume $F_k(P)$ denotes the amount of traffic on path P for routing demand k, where $P \epsilon P_k$. Thus, the total traffic routed for demand k equals $\sum_{P \epsilon P_k} F_k(P)$. On the other hand, this amount of traffic must be equal to value of demand k multiplied by the throughput, i.e. $D(k)\lambda$. This is stated by the constraint in (11).

Coding constraint: For this constraint, we need to know the amount of unicast traffic intersecting each other at $M(\xi)$. It is necessary that the traffic of input link that participate in coding, received as native, i.e. itself is not the coded traffic of another star structure. To derive this condition, we use $z_e^k(P)$ to denote the portion of the traffic on path P for demand k that is transmitted as native on link e. Thus for each combination of incoming link e_1 and outgoing link e_2 at node M, the portion of traffic that received as native is equal to $\sum_{k \epsilon D} \sum_{P \epsilon P_k : P \ni e_1 e_2} z_{e_1}^k(P)$. Since the opportunity of Star-NC arises when the relay node collects packets from all input nodes, the rate of coding at relay node is less than the rate of each incoming links. Further, as the relay node generates $N(\xi)$ packets for each group of n collected packets from input nodes, for each link-pair $e_1 e_2$ in a specific ξ, we ψ must have:

$$\frac{f^{NC}(\xi)}{N(\xi)} \leq \sum_{k \epsilon D} \sum_{P \epsilon P_k : P \ni e_1 e_2} z_{e_1}^k(P).$$

Here, $f^{NC}(\xi)$ denotes the rate by which ξ generates the coded packets and broadcasts them to the output nodes. Constraint (15) is the extension of the above constraint since the pair e_1e_2 may participate in more than one Star-NC schemes. Note that in (15), Γ refers to a set of all Star-NC opportunities in the network. We can write a balance constraint for $F_k(P)$ in terms of $z_e^k(P)$ and $f^{NC}(\xi)$ in (16) where the total transmitted traffic entering through link-pair e_1e_2 appears on LHS. The first portion on RHS, is the amount of transmitted traffic that participates in coding while the second portion is the amount that goes out as native, i.e., does not participate in any coding. Furthermore, $z_e^k(P)$ is bounded by constraints (12) and (13).

Traffic splitting constraint: Let $f(e)$ be the total flow rate of the traffic on link e. This flow for the nodes outside of any star structure is only unicast traffic. Further, as every input node of the structure, in first step of coding, sends its packet (likely XOR-ed with the overheard packets) and do not generate a coded packet without including its packet, we also consider the traffic of input links as unicast type. On the other hand, for the output links of star structures, in addition to the unicast traffic, there is NC traffic. This traffic, denoted by $f^{NC}(\xi)$, is only generated by the relay node and always broadcasted to output nodes. Thus, by the above assumption, the relay node only can transmit traffic of both unicast and NC types. Since the flow rate of NC traffic for each output links equals to $\frac{f^{NC}(\xi)}{N(\xi)}$, $f(e)$ in (17), for each output links of the star structure, will be the total traffic out of which the fraction $\frac{f^{NC}(\xi)}{N(\xi)}$ amount of it is the NC type and the rest is unicast.

Interference constraint: We consider three different link and traffic combinations: (1) the link which is not output link of any star structures and thus always transmits unicast traffic, (2) the output link of star structures which transmits NC traffic and (3) the output link of star structures with unicast traffic. This classification identifies three types of interferences which, respectively associated with three terms of constraint given in (18). In the first term, we seek link $\acute{e}$ which either resides outside of any star structure or is an input link of a star structure. In the second, we explore the star structure ξ whose NC traffic is in the interference range of link e. This holds as e is interfered with any output links of ξ. In the last term, we look for the output link $\acute{e}$ of any ξ interfering with link e whereas carries unicast traffic. The amount of unicast traffic over link $\acute{e}$ is equal to the total traffic going through $\acute{e}$ minus the rate of NC flow decoded at endpoint node of $\acute{e}$, i.e. $f(\acute{e}) - \frac{f^{NC}(\xi)}{N(\xi)}$.

Routing constraint: The constraint given by (21) maintains the flow conservation at every node of the network. For each *forwarder* node which is neither source nor destination of any session, the difference of incoming and outgoing traffic is zero. For other node which is either sink or source or both, the difference becomes a non-zero value determined by RHS of (21).

Link capacity constraint: Equations (19) and (20) limit the flow rate of a link to its capacity, for unicast and NC traffic, respectively.

Maximize λ

Subject to

$$\sum_{P\epsilon P_k} F_k(P) = D(k)\lambda \forall k \in D \tag{11}$$

$$z_e^k(P) \leq F_k(P)\ \forall k \in D, P\epsilon P_k : P \ni e \tag{12}$$

$$z_e^k(P) = F_k(P)\ \forall k \in D, P\epsilon P_k : P \ni e, t(e) = s(k) \tag{13}$$

$$\mathcal{SD}(\xi) = \{(e_1, e_2)\,|e_1 = \mathcal{S}(\xi)_i, e_2 = \mathcal{D}(\xi)_{\pi(i)} : 1 \leq i \leq n(\xi)\} \tag{14}$$

$$\sum_{\xi\in\Gamma:(e_1,e_2)\in\mathcal{SD}(\xi)} \frac{f^{NC}(\xi)}{N(\xi)} \leq \sum_{k\epsilon D}\ \sum_{P\epsilon P_k:P\ni e_1 e_2} z_{e_1}^k(P)$$
$$\forall M \in V, e_1 \in E^-(M), e_2 \in E^+(M) \tag{15}$$

$$\sum_{k\epsilon D}\ \sum_{P\epsilon P_k:P\ni e_1 e_2} F_k(P) = \sum_{\xi\in\Gamma:(e_1,e_2)\in\mathcal{SD}(\xi)} \frac{f^{NC}(\xi)}{N(\xi)}$$
$$+\sum_{k\epsilon D}\ \sum_{P\epsilon P_k:P\ni e_1 e_2} z_{e_2}^k(P) \forall M \in V, e_1 \in E^-(M), e_2 \in E^+(M) \tag{16}$$

$$f(e) = \sum_{k\epsilon D}\ \sum_{P\epsilon P_k:P\ni e} F_k(P)\ \ \forall e \in E \tag{17}$$

$$\sum_{\substack{\acute{e}\in E^C(e),\\ \exists\xi\in\Gamma^C(e):\acute{e}\in D(\xi)}} f(\acute{e}) + \sum_{\xi\in\Gamma^C(e)} f^{NC}(\xi)$$
$$+\sum_{\xi\in\Gamma^C(e)}\ \sum_{\acute{e}\in\mathcal{D}(\xi)\cap E^C(e)} \left(f(\acute{e}) - \frac{f^{NC}(\xi)}{N(\xi)}\right) \leq C(e)\ \forall e \in E \tag{18}$$

$$0 \leq f\ (e) \leq\ C(e)\ \forall e \in \mathrm{E} \tag{19}$$

$$0 \leq f^{NC}\ (\xi) \leq \min_{e\in\mathcal{S}(\xi)\cup^{\mathcal{D}}(\xi)} C(e)\ \forall\xi \in \Gamma \tag{20}$$

$$\sum_{e\epsilon E^+(v)} f\ (e) - \sum_{e\epsilon E^-(v)} f\ (e)$$
$$= \begin{cases} 0 & \forall v \neq s\ (k)\ , d\ (k)\ , k \in D \\ \sum_{\substack{k\in D,\\ v=s(k)}} \lambda_k - \sum_{\substack{k\in D,\\ v=d(k)}} \lambda_k & \forall v = s\ (k)\ , d\ (k)\ , k \in D \end{cases} \tag{21}$$

5 Evaluation

In this section, we compute the performance of the above joint routing and NC scheme with the non-NC schemes. First we introduce the configurations including the network topologies, traffic models, coding schemes and routing strategies. Next, we introduce our self-designed testbed tool for evaluation. Finally, the evaluation results are presented.

5.1 System Configurations

5.1.1 Network Topologies

The first target topology used for evaluation is shown in Fig. 6. The network consists of 49 nodes in a square of side 7 × 7 units where the positions of the nodes were chosen randomly while maintaining connectivity. The second topology is a 7 × 7 grid network. For both topologies, we assume that the transmission and interference range to be 1.7 unit and 2 units, respectively. The average node degree for random and grid network are, respectively equal to 6.0 and 6.4. We employ the Star-NC scheme for the star structures of size 2 and 3. Since the pattern leading to a full-star structure of size 2 is the same as the COPE "X" topology [7], we exclude this coding from the Star-NC scheme and accounts it as the COPE-type coding scheme

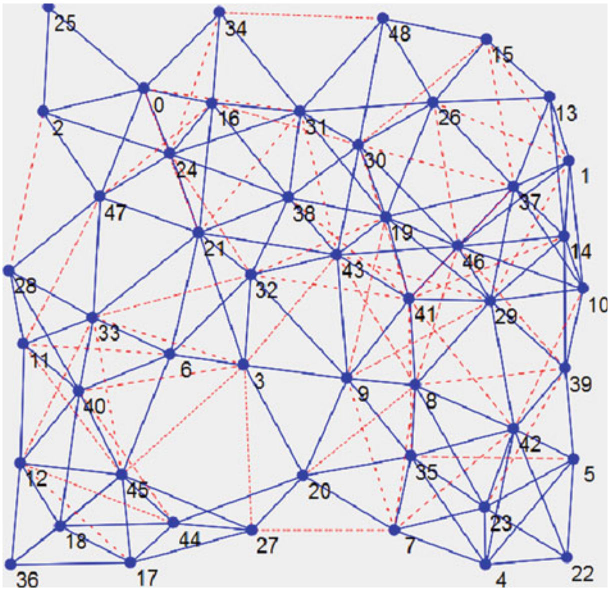

Fig. 6 The mesh network topology used for evaluation consisting of 49 nodes randomly located in a square of size 7 × 7. The communication/interference range are equals to 1.7 and 2. The *solid/dash lines* illusttrate the node-pairs which are in communication/interference range of each other

5.1.2 Routing Strategies

We consider three routing strategies: (1) single-path routing (SP), (2) optimized single path routing (OSP) and (3) multi-path routing (MP). In particular, the single-path routing can be obtained by Dijkstra's algorithm and a metric which is to be minimized such as the Hop-count, joint Hop-count and physical distance or ETX metric [20]. Note that the single-path routing is neither coding-aware nor interference-aware, i.e. the paths are selected based on the desired metric and thus neither coding opportunities nor interference among the nodes is considered. To overcome these shortcomings, we are lead to use multi-path routing.

As mentioned above, our LP formulation supports both the multi-path and single-path routing. Multi-path routing considers *interference-aware* routing where the paths minimizing interference are selected. We implement multi-path routing by ideal of *internally pairwise edge-disjoint paths* [21]. That is, for two nodes s and d, first we find the shortest single path between s and d. Then we remove links of the current path and explore other possible path (i.e., shortest- single) among the remaining links. We repeat the procedure until no route exists between s and d. At the end, we remove the possible cycles from the paths found between $s$and d in presence of all the removed links. The number of the paths found this method depends on the *edge-connectivity* of s and d in the graph, which equals to the *edge-cut* of s and d [21]. In our evaluation for the full grid network, this number is a variable between 3 and 8 depending on the position of the source and destination nodes.

However, in practice, multi-path strategy is not applicable in most cases due to the high routing maintenance overhead. Thus, similar to [13], we consider *optimized single-path* routing in addition to single-path and multi-path routing. The key idea of the optimized single-path routing is to select the path that provides the maximum flow among the paths selected by multi-path routing for each session. In our evaluation, the optimized single-path routing is obtained by the following steps. First, we solve the LP with multi-path routing. For each session, we select the path that achieves the highest flow, i.e., $P_{Opt} = Max_{P \in P_k} F_k(P)$. Second, the chosen paths are then fed to the LP and solve the LP again.

Note that optimized single-path routing like multi-path routing tries to minimize interference among nodes, i.e. interference-aware routing. Intuitively, multi-path routing can provide more coding opportunities for the wireless NC than single-path routing. Hence, when the NC scheme is employed in either multi-path or optimized single-path routing method, *coding awareness* is added to optimization options. In this case, the routing protocol tries to transmit traffic from the paths which create more coding opportunities in addition to minimize the interference among nodes. Hence, for multi-path routing, both the notion of coding-aware and interference-aware routing must be considered. In particular, our LP formulation provides a systematic approach for finding the routes that optimize the tradeoffs between the opposite effects of increased coding and increased interference and identifies the best routing choices.

5.1.3 Coding Strategies

Our evaluation covers three main NC schemes 1) the Star-NC with partial star 2) the Star-NC scheme with both partial and full star 3) the COPE-type NC scheme which are referred to as NC(STAR.P) , NC(STAR) and NC(COPE) in the plots, respectively. COPE-type NC evaluations are based on the LP formulations in [8]. Since we only use the star structures of size 2 and 3, the radius of overhearing is equal to 1. Thus STAR.P refers to two-hop coding scheme. Further, STAR contains full Star-NC of size 3 in addition to STAR. Moreover, we consider two scenarios for joint Star and COPE type NC indicated by NC(STAR.P+COPE) and NC(STAR+COPE) in the plots. The former considers only partial star structure while the latter considers both partial and full ones. For both schemes, we extend our LP formulation to covers COPE-type NC in addition to the Star-NC scheme.

5.1.4 Traffic Models

We focus on two different random traffic models named as *random* and *directional*. For the random model, as the name implies, the source and destination of each session are chosen randomly from the network nodes. However, for directional model, the nodes are partitioned into two consecutive groups named as *senders* and *receivers*. For each flow, the source and destination are picked up from the sender and receiver groups, respectively. We use this model since it suits practice for wireless mesh networks. For both models, we vary the number of demands from 50 to 500.

5.2 Putting it all Together

We developed a testbed tool which integrates all of the above modeling options. By setting the proper configurations about network topology, traffic model, coding scheme and routing strategy, our evaluation testbed generates the corresponding LP system. We solve this LP using AMPL [22] with the CPLEX solver [23] to obtain the theoretically optimized throughputs and the corresponding flows for the non-NC, Star-NC, COPE-type NC and joint Star and COPE NC schemes, respectively. For most often cases of multi-path routing, it is necessary to solve the LP formulation again. Indeed, the load of some links becomes zero while the throughput is evaluated by considering corresponding interference constraints about these links. Thus, after each LP solution for multi-path routing, our testbed tool verifies whether

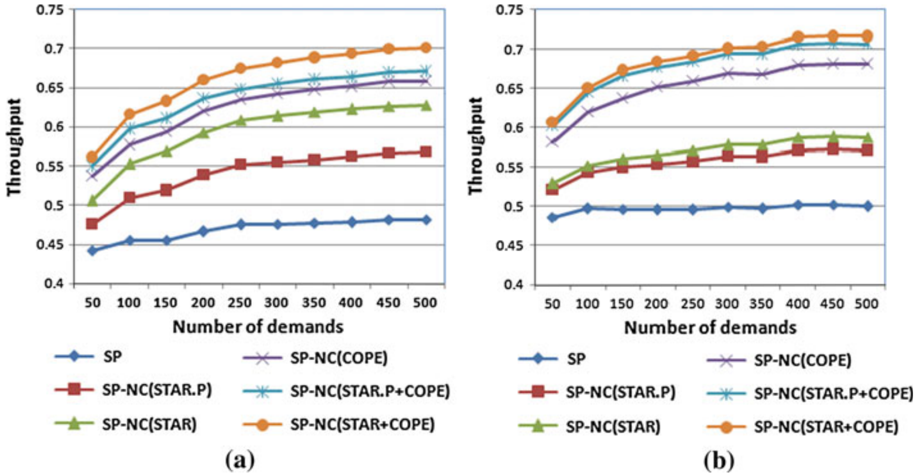

Fig. 7 Network throughput (normalized to link capacity) for random traffic model. **a** For grid network. **b** For random network

zero-load links are presented, if yes, then it removes the paths passed through these links and solves the LP system again.

We assume an ideal MAC layer with both the lossless links and an optimal medium access algorithm, i.e., the channel is fairly assigned to nodes that have a packet for sending. For simplicity, we assume all channels have the same capacity. Obviously the ideal MAC assumption is held for both coding and non-coding schemes, i.e. we evaluated the performance of non-coding scheme under the assumption of ideal MAC. In COPE, as a first testbed deployment of NC, reported that the throughput improvements for a real MAC (802.11 family) is higher than the theoretical value in some situations such as UDP transmissions. This additional gain is basically originated from intrinsic unfairness of 802.11 MAC. Using a careful model for an 802.11 like MAC within our theoretical framework is more challenging and will be the subject of our future work.

5.3 Evaluation Results

Evaluation 1: Coding for random traffic in grid network: Fig. 7a shows the benefits of coding for single-path routing. Since the routes in single-path are fixed, the throughput improvement is basically originated from the coding opportunities created by the unicast sessions crossed at a relay node. The improvement for STAR.P, STAR and COPE are approximately equal to 18, 30 and 35 % relative to non-coding scheme. Further joint star and COPE coding has performance improvement about 39 and 45 % relative to single-path routing, respectively. It means that Star-NC creates coding opportunities different from COPE-type coding schemes which can improve the gain of coding up to 10 % relative to COPE-Type NC schemes.

Evaluation 2: Coding for random traffic in random network: Fig. 7b repeats the previous evaluation in random network. The improvement for STAR.P, STAR and COPE are approximately equal to 15, 18 and 35 % relative to non-coding scheme. We can see a significant reduction in Star-NC opportunities in random network relative to grid network. Further both of the joint coding schemes have a performance improvement about 43 % relative

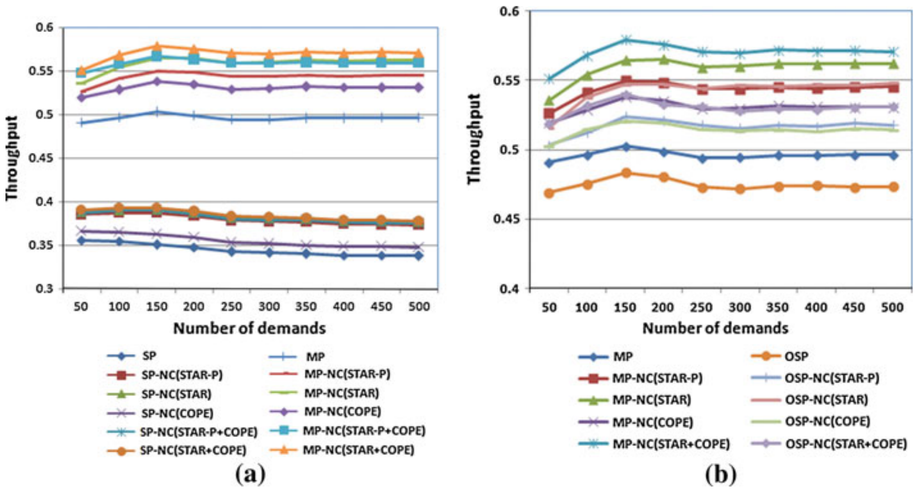

Fig. 8 Network throughput (normalized to link capacity) for directional traffic model. **a** Single path versus multi-path. **b** Multi-path versus optimized single path

to single-path routing. Thus in spite of reduction in Star-NC opportunities, it improves the coding gain up to 7 % relative to COPE-Type NC schemes.

Evaluation 3: Coding for directional traffic in random network: We repeat the previous evaluation for a directional traffic model. We divide the 7×7 square nodes into two sub squares of size 3.5×7 named as *left* and *right* group. Each flow is started from a node in left group and ended to a one in right group. As a first observation, a significant gap exists between the aggregated throughputs of this model and the corresponding results for the random traffic model. As shown in Fig. 8, the throughput varies from 0.3 to 0.4 depending on the routing-coding choices while throughput range for the previous model was between 0.45 and 0.75.

According to Fig. 8a, the improvement for COPE is very limited while the gain for Star-NC is significant. Precisely, COPE gain is limited to 3 % while both variations of Star-NC have a gain about 12 %. Indeed, Star-NC outperforms the COPE-type NC schemes for directional traffic. It is because the coding opportunities for some COPE-type schemes such as Alice-Bob are not created due to the pattern of traffic and most of the opportunities are limited to the "X" topologies. Moreover, we observe that the joint Star and COPE-type coding has no significant benefit over Star-NC alone.

Note that, this model resembles the typical traffic on wireless mesh networks that provide Internet connectivity for end users. The gateway nodes connect the end user machines to Internet. Since a wide range of network applications have client/server models where most of users act as client, a large portion of the traffic is related to the flows from the gateway nodes to the client nodes.

Evaluation 4: Coding for multi-path routing: Multi-path routing selects the paths which minimize the interference among nodes. Since the average node degree in our network topology is relatively high (≈ 6.0), MP always finds multiple paths between each pair of nodes. Figure 8a also shows the results for multi-path routing. As shown in the figure, MP itself increases throughput by 45 % relative to SP. The coding for MP, COPE, STAR.P and STAR bring approximate gains of 10, 15 and 20 % relative to MP. Further, the results show that the employing the joint Star and COPE has a slight performance improvement about 5 % over COPE alone. Note that the coding gain for the joint of STAR.P and COPE are about 20 %

which is equal to gain of STAR alone. Moreover the gain of joint STAR and COPE is 23 % which shows a slight performance (≈3 %) over STAR alone.

Evaluation 5: Coding for optimized single-path routing: As mentioned earlier, the optimized single path, instead of using shortest cost, tries to find the paths that reduce the total interference. By adding the option of coding, it becomes a coding-aware routing in addition to interference-aware feature. Figure 8b shows the results for optimized-single path versus multi-path routing. We can see that OSP itself increases the throughput about 40 % relative to SP. Note that, this improvement is near to MP for the directional traffic model. Applying NC opportunities to OSP provides additional gain about 12, 13, 22 and 25 %, respectively for COPE, STAR.P, STAR and STAR+COPE. Briefly, the results show that the difference between throughput performance of MP and OSP is at most 7 % for all version of coding.

6 Application to Existing Wireless Mesh Networks

Currently, there has been a growing interest in deploying Wireless Mesh Network by municipalities. It basically originates from the desire to provide broadband connectivity to mobile users. These networks create more opportunities for new wireless-base application and service and thus generate revenue to local business. Now, Cities in US such as San Francisco, Mountain View, Philadelphia and Boston have initiated or finished projects with the aim to provide wireless broadband service to their citizens [24]. These implementations could be taken as real model of wireless mesh network which we try to redesign them with considering the intrinsic characteristics of wireless network such as broadcast media nature and data redundancy. Next, we give a brief description of Mountain View project. This project which was done by Google, made Mountain View the first city in the Bay Area to get a full umbrella of free WiFi coverage for free wireless Internet access.

It started with the idea of providing broadband connectivity to the police department especially to the officer driving in duty. Then, it was expanded to include the whole town. The project is based on IEEE 802.11 b/g. The hierarchical wireless mesh network architecture starts with Tropos single radio 802.11 b/g mounted on city owned light poles. These nodes act as the Access Points (APs) and use the same frequencies to cover different areas of the city. Google installed about 460 APs in whole town, which translates into 30 to 40 APs per sq. mile. The mesh network consists of each of these APs wirelessly connected to a designated Gateway. There is one gateway for every six APs (5–7 gateways per square mile). Each deployed Base Station covers 4 square mile, namely, three Base Stations are needed to cover the entire city. Base stations and gateways are placed on rooftops (high riser school). The project was completed in August 2006 and it has been serving the citizens of that area.

7 Discussion and Conclusion

We proposed Star-NC, a network coding scheme over star structures for multiple parallel unicast sessions. The key idea of Star-NC is the flexibility of the opportunistic listening which creates more coding opportunities. We provided an analysis of our proposed NC scheme under realistic assumptions. We gave bounds on the gain as well as encoding matrices for various routing patterns of unicast transmissions. We also provided a linear programming framework to investigate and optimize the network throughput of the proposed coding in wireless mesh networks. Our formulation is general; hence it works for every wireless topology and every routing strategy. Further, we integrated other coding scheme such as COPE and integrated them in our formulation to compare the performance of the Star-NC scheme with other NC schemes.

Our analysis and evaluations confirm that the combination of our network coding scheme with different routing strategies is capable of achieving substantial improvement in overall network throughput compared to the non-coding schemes. More specifically, we showed that for the random traffic model, the performance of Star-NC is between 18 and 35 % depends on the network topology. However, for the directional traffic, the Star-NC has a significant benefit over the COPE-type NC scheme. In addition, we studied the employment of joint Star and COPE coding opportunities in a new extensive scheme. The results showed that the joint scheme has higher performance that each alone. In summary, the performance of the COPE-type NC scheme can be improved by exploiting the coding opportunities created by Star-NC.

Appendix A: Proof for Theorem 1

We explore a sufficient condition for Theorem 1 by proper selection of $\mathbf{Q}_{\boldsymbol{M}}$ and $\mathbf{Q}_{\boldsymbol{B}}$ in which every output nodes is capable of decoding its packet. The output node D_k has a system of η linear equations, out of which $\eta - m$ of them are received from the relay node's encoded packets ($\mathbf{Q}_{\boldsymbol{M}}$) and the others are captured by opportunistic listening of neighbor nodes. Therefore, it suffices to show that each output node has enough information to solve its system and decode all of the packets. This condition, for $k > m$, represents a feature of $\mathbf{Q}_{\boldsymbol{M}}$ in which the deletion of the columns i_{k-m}through i_{k-1} results a non-zero determinant matrix. However, for $k \leq m$ this deletion is occurred for the columns $i_1$to i_k of the union of $\mathbf{Q}_{\boldsymbol{M}}$ and a matrix formed by the first $m - k$ rows of $\mathbf{Q}_{\boldsymbol{B}}$, and must lead to a non-zero determinant matrix. Next, we use this condition and try to present a constructive proof for Theorem 1 by seeking suitable $\mathbf{Q}_{\boldsymbol{M}}$ and $\mathbf{Q}_{\boldsymbol{B}}$.

Case $\boldsymbol{m} = \mathbf{1}$: Consider $\mathbf{Q}_{\boldsymbol{M}}$ and $\mathbf{Q}_{\boldsymbol{B}}$ matrices as:

$$\mathbf{Q}_{\boldsymbol{M}} = \begin{pmatrix} \mathbf{1} & 0 & \cdots & 0 & 1 \\ 0 & \mathbf{1} & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & 1 \end{pmatrix}_{(\eta-1)\times\eta} \qquad \mathbf{Q}_B = I_\eta \tag{22}$$

For any η (even for $\eta > 3m + 2$), deletion of each column of $\mathbf{Q}_{\boldsymbol{M}}$ produces a non-zero determinant rectangular matrix.

Case $\boldsymbol{m} = \mathbf{2}$: For $\eta \geq 4$, Consider $\mathbf{Q}_{\boldsymbol{M}}$ as concatenation of the identity matrix of size $\eta-2$ and a two-columns matrix J, i.e. $\mathbf{Q}_{\boldsymbol{M}} = \left(\mathbf{I}_{(\eta-2)}|\mathbf{J}_{(\eta-2)\times 2}\right)$ where, for different values of η, J is defined as in Table 3. Now, we seek for the pairs of $\mathbf{Q}_{\boldsymbol{M}}$ columns whose deletion produces a zero determinant matrix. This happens in two different cases as either the generation a row of zero elements or remaining a set of linearly dependent columns. Note that, since the first part of $\mathbf{Q}_{\boldsymbol{M}}$ is an identity matrix, it has no linearly dependent rows and thus the deletion a pair of columns (except for η =4) does not leave a matrix with dependent rows. For simple verification, we define tow terms $ZR(\mathbf{Q}_{\boldsymbol{M}})$ and $CD(\mathbf{Q}_{\boldsymbol{M}})$ to denote the sets containing the tuples of the columns of $\mathbf{Q}_{\boldsymbol{M}}$. In the former, each tuple refers to the columns of $\mathbf{Q}_{\boldsymbol{M}}$ whose deletion creates a row of zeros, while in the latter, it refers to the columns whose sum is a zero vector. For $ZR(\mathbf{Q}_{\boldsymbol{M}})$, we consider in tuples with at most two elements while for $CD(\mathbf{Q}_{\boldsymbol{M}})$ the tuples with at most $\eta-2$ elements are concerned. By a little effort, we can compute these

Table 3 $J_{(\eta-2)\times\eta}$, $ZR(Q_M)$ and $CD(Q_M)$ for $m = 2$

η	$J_{(\eta-2)\times\eta}$	$\mathbf{ZR(Q_M)}$	$\mathbf{CD(Q_M)}$
4	$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$	(1, 4), (2, 4)	(3), (1, 2, 4)
5	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$	(1, 5), (2, 4)	(1, 3, 5), (2, 3, 4), (**1**, **2**, **4**, **5**)
6	$\begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$	(1, 6), (2, 6), (3, 5), (4, 5)	(1, 2, 6), (3, 4, 5)
7	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$	(1, 7), (2, 6)	(**1**, **2**, **6**, **7**), (1, 3, 4, 5, 7), (2, 3, 4, 5, 6)
8	$\begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$	(1, 8), (2, 8), (3, 7), (4, 7)	(1, 2, 5, 6, 8), (3, 4, 5, 6, 7), (**1**, **2**, **3**, **4**, **7**, **8**)

sets, shown in Table 3, for the above $\mathbf{Q_M}$. Let $D_\eta^2(Q_M)$ denotes a set of column pairs of $\mathbf{Q_M}$ whose deletion results a zero determinant matrix. We can write:

$$D_\eta^2(Q_M) = ZR(Q_M) \cup \{(i, j) \,|\exists x \in CD(Q_M) : x \subseteq (1, 2, \ldots, \eta) - (i, j)\} \tag{23}$$

Now, we try to find a relation between the ability of decoding for each target permutation and $D_\eta^2(Q_M)$. We define a target permutation as $\mathrm{Q_M}$-proper if its first $\eta - 1$ elements do not include any pair of $D_\eta^2(Q_M)$ as two consecutive elements. This condition guaranties the feasibility of decoding for output nodes D_3 through D_η. Obviously, only a subset of the permutations will be $\mathrm{Q_M}$-proper. For each the others, we try to exchange the columns of $\mathbf{Q_M}$ in a way that it becomes a proper permutation for the altered $\mathbf{Q_M}$. The main idea is that we exchange of the columns based on the mapping which converts a non-proper permutation to a proper one. For instance, let $\rho = (1, 2, 7, 6, 8, 5, 4, 3)$ and $\pi = (6, 5, 1, 2, 8, 3, 7, 4)$, it is simple to verify that ρ is $\mathrm{Q_M}$-proper, but π is not. Now, we can create a new matrix $\acute{\mathbf{Q}}_M$ from $\mathbf{Q_M}$ by the following exchanges:

$$X_{\rho\to\pi} = (1 \leftrightarrow 6, 2 \leftrightarrow 5, 7 \leftrightarrow 1, 6 \leftrightarrow 2, 8 \leftrightarrow 8, 5 \leftrightarrow 3, 4 \leftrightarrow 7, 3 \leftrightarrow 4)$$

Now, π is a $\acute{\mathbf{Q}}_M$-proper permutation. We refer to ρ as *generator* of π. It seems that a single $\mathrm{Q_M}$-proper generator, like ρ, is enough for all target permutation. However, we have not verified the ability of decoding for the output nodes D_1 and D_2 yet. Indeed, these nodes overhear the input nodes S_1 and S_2 and thus decoding process is further dependent on $\mathbf{Q_B}$. We fix our choices of $\mathbf{Q_B}$ to an upper triangular matrix with elements of 1 in the main diagonal and all rows expect the second one. This $\mathbf{Q_B}$ implies that every input node S_i $(i \neq 2)$ sends the encoded packet $\sum_{k=i}^{\eta} P_k$ to the relay node, i.e., S_i needs to overhear the transmissions of S_{i+1} and S_{i+2}.

On the other hand, since both D_1 and D_2 can overhear transmission of S_1, i.e., the encoded packet $\sum_{i=1}^{\eta} P_i$. In addition to this packet, D_1 overhears the native packets P_2 and D_2

overhears $P_{\pi(1)}$. According to the overheard information, the decoding of D_1 and D_2 reaches to system of equations whose core is a matrix resulted from $\mathbf{Q_M}$ (or $\acute{\mathbf{Q}}_{\mathbf{M}}$ by adding a row of 1 and deleting the column either 2 (for D_1) or $\pi(1)$ (for D_1). Thus, it is enough to consider the effect of a column deletion in to account either for creation of a row with zero elements or for remaining a set of dependent columns. Since $ZR(Q_M)$ contains only the tuples of size 2, zeroing a row is not taken by a single column deletion. Further, since the core matrix has a row full of 1, generation a set of dependent column only is occurred for the even tuples of $CD(Q_M)$. As for $\eta = 4$ and $\eta = 6$, $CD(Q_M)$ has no even size tuples, the resulting matrices do not have any dependent columns, and hence is suitable for any target permutation. But for $\eta = 5$, 7 and 8, the resulting matrices have inherited the even size column dependencies of $\mathbf{Q_M}$. Thus, the decoding process could be done if the deletion is occurred for a column among these tuples.

We satisfy this condition by selecting a suitable $\mathrm{Q_M}$-proper generator ρ for each π in a manner that both the columns mapped to 2 and $\pi[1]$ via $X_{\rho\rightarrow\pi}$, are picked up from these tuples. By selecting ρ such that $\rho[1] = 1$, the condition related to column $\pi[1]$ is met since the first column is appeared in of all the mentioned tuples. The other condition, corresponding to columns 2, is satisfied by choosing a ρ that the elements of each dependent tuple are located in even positions of ρ when the position of 2 in π is even and vice versa when the position of 2 in π is odd. Thus, one of the following options is suitable for the generator permutation:

Position of 2 in π	$\eta = 5$	$\eta = 7$	$\eta = 8$
odd	(**1**, 4, **5**, 3, **2**)	(**1**, 3, **2**, 4, **6**, 5, 7)	(**1**, 2, **3**, 4, **8**, 6, **7**, 5)
Even	(1, **4**, 5, **2**, 3)	(1, **2**, 3, **6**, 4, **7**, 5)	(1, **2**, 3, **4**, 6, **7**, 5, **8**)

Lastly, for m = 3 and m = 4, we used an extension of the method for m = 2 with a subtle difference that the most of verification steps is done by a computerized program. This is due to largeness of state space of the problem. The details of proof are deleted for the lack of space.

References

1. Ahlswede, R., Cai, N., Li, S.-Y. R., & Yeung, R. W. (2000). Network information flow. *IEEE Transactions on Information Theory*, *46*(1), 1204–1216.
2. Li, S. Y., Yeung, R. W., & Cai, N. (2003). Linear network coding. *IEEE Transaction on Information Theory*, 49(2), 371–381.
3. Li, Z., & Li, B. (2004). Network coding in undirected networks. In *Proceedings of the CISS*.
4. Li, Z., & Li, B. (2004). Network coding: The case for multiple unicast sessions. In *Proceedings of the Allerton conference on communications*.
5. Wu, Y., Chou, P. A., & Kung, S. Y. (2004). *Information exchange in wireless networks with network coding and physical-layer broadcast*. MSR-TR-78.
6. Koetter, R., & Ho, T. (2005). Online incremental network coding for multiple unicasts. In *Proceedings of the DIMACS working group on network coding*.
7. Katti, S., Rahul, H., Hu, W., Katabi, D., Medard, M., & Crowcroft, J. (2008). XOR in the air: Practical wireless network coding. *IEEE/ACM Transactions on Networking*, *16*, 497–510.
8. Sengupta, S., Rayanchu, S., & Banerjee, S. (2007). An analysis of wireless network coding for unicast sessions: The case for coding-aware routing. In *Proceedings of the IEEE INFOCOM*.
9. Omiwade, S., Zheng, R., & Hua, C. (Jan. 2008). Butterflies in the mesh: lightweight localized wireless network coding. In *Proceedings of NetCod*, Lausanne, Switzerland.

10. Wang, C.-C., & Shroff, N. B. (2010). Pairwise intersession network coding on directed networks. *IEEE Transactions on Information Theory, 56*(8), 3879–3900.
11. Yuan, J., Li, Z., Yu, W., & Li, B. (2006). A Cross-layer optimization framework for multihop multicast in wireless mesh networks. *IEEE Journal on Selected Areas in Communications,* 24(11) 2092–2103.
12. Zhang, X., & Li, B. (2008). On the benefits of network coding in multi-channel wireless networks. In *Proceedings of the IEEE INFOCOM*.
13. Zhang, H., & Su, X. (June 2009). Modeling throughput gain of network coding in multi-channel multi-radio wireless ad hoc networks. *International Journal of Selected Area in Communications, 27*(5).
14. Liu, J., Coeckel, D., & Towsley, D. (2007) Bounds on the gain of network coding and broadcasting in wireless networks. In *Proceedings of the IEEE INFOCOM*.
15. Gupta, P., & Kumar, P. R. (2000). The capacity of wireless networks. *IEEE Transactions on Information Theory, 46*(2), 388–404.
16. Le, J., Lui, J. C. S., & Chiu, D. M. (2008). How many packets can we encode? An analysis of practical wireless network coding. In *Proceedings of the IEEE INFOCOM*.
17. Kwon, S., Hendessi, F., & Fekri, F. (2009). Cooperative network coding and coding-aware channel assignment in multi-channel. Multi-interface wireless networks. In *Proceedings of the IEEE SECON*.
18. Kwon, S. C., Hendessi, F., Fekri, F., & Stüber, G. L. (Jan 2011). A novel collaboration scheme for multi-channel/interface network coding. *IEEE Transactions on Wireless Communications, 10*(1), 188–198.
19. Fragouli, C., Katabi, D., Markopoulou, A., Medard, M., & Rahul, H. (2007). Wireless network coding: Opportunities and challenges. In *MILCOM*.
20. De Couto, D. S. J., Aguayo, D., Bicket, J., & Morris, R. (2003). A high-throughput path metric for multi-hop wireless routing. In *Proceedings of the ACM MOBICOM* (pp. 134–146).
21. West, D. B. (2001). *Introduction to graph theory* (2nd ed., pp. 150–170). s.l.: Prentice Hall.
22. AMPL: A modeling language for mathematical programming. http://www.ampl.com/. (Online) IBM Inc.
23. CPLEX, ILOG. http://www.ilog.com. (Online) ILOG Inc.
24. Hassan-Ali, M. (2007). *Municipal wireless mesh networks as a comptetive*. broadband delivery platform M.S. Thesis, Massachusetts Institute of Technology.

Author Biographies

Alireza Shafieinejad got the B.S degree in computer engineering form Sharif University of Technology, Tehran, iran, in 1999. He is also received M.Sc. and Ph.D. degree in computer engineering from Isfahan University of Technology, Isfahan, Iran, respectively, in 2002 and 2012. His research interests are in the area of Wireless Network Coding, Cross-Layer Optimization and Block Cipher Design and Analysis.

Faramarz Hendessi received a B.Sc. degree from Baluchestan University, Iran in 1986, and an M.Sc. degree from Isfahan University of Technology, Iran in 1988, both in Electrical Engineering. In 1993 he received a Ph.D. in Electrical Engineering from Carleton University, Ottawa, Ontario, Canada. He is currently with the Department of Electrical and Computer Engineering at Isfahan University of Technology. Dr. Hendess's research interests are in Wireless Network, Signal Processing and Network Security.

Faramarz Fekri received the B.Sc. and M.Sc. degrees from Sharif University of Technology, Tehran, Iran, in 1990 and 1993, respectively, and the Ph.D. degree from the Georgia Institute of Technology in 2000. Since 2000, Dr. Fekri has been with the faculty of the School of Electrical and Computer Engineering at the Georgia Institute of Technology. He is a member of the Georgia Tech Broadband Institute (GTBI), the Center for Signal and Image Processing (CSIP), and the Georgia Tech Information Security Center Dr. Fekri's current research interests are in the area of communications and signal processing, in particular error-control coding, information processing for wireless and sensor networks, and communication security. He has introduced finite-field wavelets and their applications to coding and cryptography. At Georgia Tech, he leads research in reliable and secure content delivery in wireless ad hoc (sensor) networks, space-time and delay tolerant networks.